Вісник Київського університет 2010, 2 *Bulletin of University of Kyiv*
Серія: фізико-математичні науки *Series: Physics & Mathematics*

УДК 519.8

Горбачук В.М., к.ф.-м.н., с.н.с., Гаркуша* Н.І., к.е.н.

Необхідні умови першого порядку для задач дворівневої оптимізації

Для досить гладких задач дворівневої оптимізації з обмеженнями виведені необхідні умови першого порядку, виходячи з результатів параметричного програмування та Кларка. Наведено приклад такої задачі при проектуванні мережі.

Ключові слова: дворівнева оптимізація, параметричне програмування, Кларк, мережа.

V. M. Gorbachuk, Ph.D. (Physics and Mathematics), N.I. Garkusha*, Ph.D. (Economics)

First order necessary conditions for two-level optimization problems

Based upon the results of parametric programming and those of Clarke, the first order conditions have been derived for suffciently smooth two-level constrained optimization problems. The example of such problem is given for network design.

Key words: two-level optimization, parametric programming, Clarke, network.

*E-mail: GorbachukWasyl@netscape.net, NGarkusha@gmail.com
Статтю представив проф., док. ф.-м.н. Хусаїнов Д.Я.

Розглянемо нелінійну задачу проектування мережі [1]. Нехай середні витрати (cost) на ланці (link) a із загальним потоком (flow) $f(a,x)$ рівні

$$C_a[x_a, g(a,x)] = t_a\left\{1+\rho\left[\frac{g(a,x)}{s_a+y_a}\right]^\sigma\right\},$$

де: s_a – наявна місткість ланки a; x_a – приріст місткості ланки a; x – загальний вектор $\{x_a\}$; t_a, ρ, σ – деякі параметри. Тому загальна сума суспільних витрат на транспортування становить

$$S = \sum_a \int_0^{g(a,x)} M_a(x,y)\,d\,y,$$

де граничні витрати на ланці a з потоком $g(a,x)$ визначаються

$$M_a[x_a, g(a,x)] = \frac{d\{C_a[x_a, g(a,x)]\times g(a,x)\}}{d[g(a,x)]}.$$

Припустимо, сумарні інвестиційні витрати для поліпшення транспортних ланок становлять

$$I = \sum_a G_a(x_a),$$

де $G_a(x_a)$ – інвестиційна функція від збільшення місткості на x_a ланки a. Тоді на верхньому рівні державний плановик обирає таке значення вектора x, що мінімізує суму витрат

$$F = S + \theta I \to \min_x \qquad (1)$$

за обмеження

$$x_a \geq 0, \qquad (2)$$

де θ – відносна вага інвестиційних витрат. На нижньому рівні споживачі вибирають рівноважні за Уордропом [2] маршрути: y є розв'язком підзадачі

$$f = \int_0^{g(a,x)} C_a(x,y)\,d\,x \to \min_y \qquad (3)$$

при обмеженнях

$$\sum_p Y_{ijp} = T_{ij} \quad \forall i,j, \qquad (4)$$

$$Y_{ijp} \geq 0 \quad \forall p,i,j, \qquad (5)$$

$$g(a,x) = \sum_i \sum_j \sum_p Y_{ijp}\delta_{ijp}^a \quad \forall a, \qquad (6)$$

де: Y_{ijp} – потік від i до j на шляху (path) p; δ_{ijp}^a – булева змінна, що рівна 1 тоді й тільки тоді, коли ланка a належить такому шляху.

Задача (1)–(6) є прикладом задачі дворівневої оптимізації.

Теорема 1. Нехай x^* – це точка локального інфімума по $x \in A$ неперервно диференційованої функції $F(x,y): R^n \times R^m \to R$, де:

y – це точка мінімума по $z \in \Omega(x)$ двічі неперервно диференційованої функції $f(x,z): R^n \times R^m \to R$;

A – компактна підмножина в R^n, яка містить відкриту обмежену підмножину B;

$\Omega(x)$ – це підмножина точок $z \in R^m$, які задовольняють рівностям $\Phi_i(x,z)=0$, $i=1,...,p$, та нерівностям $\Psi_j(x,z)\le 0$, $j=1,...,s$;

$\Phi_i(x,z)$ – двічі неперервно диференційована функція, аффінна по z;

$\Psi_j(x,z)$ – двічі неперервно диференційована функція, опукла по z.

Позначимо $S(x)$ множину всіх таких точок y.

1) Припустимо, $\forall y \in S(x)$, $\forall x \in B$ лінійно незалежними є градієнти $\nabla_y \Phi_i(x,y)$, $i=1,...,p$, та градієнти $\nabla_y \Psi_k(x,y)$ на активних обмеженнях $k \in I(x,y)=\{j\in\{1,...,s\}:\Psi_j(x,y)=0\}$.

2) Стандартна функція Лагранжа (Lagrange)

$$L(x,y,\mu,\lambda)=f+\sum_{i=1}^{p}\mu_i\Phi_i+\sum_{j=1}^{s}\lambda_j\Psi_j$$

задовольняє сильним достатнім умовам другого порядку: якщо $x_0 \in B$, $y_0 \in S(x_0)$, $\vec{0}\ne h \in R^m$,

$\left\langle \nabla_y \Phi_i(x_0,y_0),h\right\rangle=0$, $i=1,...,p$,

$\left\langle \nabla_y \Psi_l(x_0,y_0),h\right\rangle=0$ на підмножині обмежень $l \in J(x_0,y_0)=\{l\in I(x_0,y_0):\lambda_{l0}>0\}$, то

$$\left\langle h,Q(x_0)\right\rangle>0,$$

де:

$$Q(x_0)=\nabla^2_{yy}L(x_0,y_0,\mu_0,\lambda_0);$$

за припущенням 1), вектори множників $\mu_0 \in R^p$, $\lambda_0 \in R^s_+$ (R^S_+ – це невід'ємний ортант R^s) є єдиними, що задовольняють умовам Куна–Такера (Kuhn–Tucker):

$$\nabla_y L(x_0,y_0,\mu_0,\lambda_0)=0, \quad (7)$$

$$\lambda_{j0}\Psi_i(x_0,y_0)=0,\ j=1,...,s. \quad (8)$$

3) Нехай t – це єдиний розв'язок приєднаної (adjoint) задачі квадратичного програмування пошуку інфімума по t функції

$$\frac{1}{2}\left\langle t,Q(x^*)t\right\rangle-\left\langle \nabla_y F(x^*,y^*),t\right\rangle \quad (9)$$

за обмеження

$$C(x^*)t=0, \quad (10)$$

де: $y^*=S(x^*)$;

$I(x^*,y^*)=J(x^*,y^*)$;

$C(x^*)$ – матриця, рядками якої є частинні градієнти $\nabla_y\Phi_i(x^*,y^*)$, $i=1,...,p$, та $\nabla_y\Psi_k(x^*,y^*)$, $k\in I(x^*,S(x^*))$, $y^*=S(x^*)$.

Позначимо η єдиний множник Лагранжа задачі (9)–(10). Тоді

$$\vec{0}\in\nabla_x F(x^*,y^*)-[\nabla^2_{yx}L(x^*,y^*,\mu^*,\lambda^*)]^T t-$$
$$-[D(x^*)]^T\eta+N_A(x^*),$$

де: μ^*, λ^* – вектори множників, які задовольняють умовам (7)–(8) при $x_0=x^*$, $y_0=y^*$, $\mu_0=\mu^*$, $\lambda_0=\lambda^*$;

$D(x^*)$ – матриця, рядками якої є частинні градієнти $\nabla_x\Phi_i(x^*,y^*)$, $i=1,...,p$, та $\nabla_x\Psi_k(x^*,y^*)$, $k\in I(x^*,S(x^*))$, $y^*=S(x^*)$;

$N_A(x^*)=\{v^*\in R^n:\left\langle v,v^*\right\rangle\le 0 \quad \forall v\in C_A(x^*)\}$ – це нормальний конус Кларка (Clarke), або від'ємний полярний конус до $C_A(x^*)$;

$C_A(x)=\left\{v\in R^n:\lim\limits_{\lambda\downarrow 0;u\to x;u\in A} dist\left(\frac{u+\lambda v}{\lambda}\right)=0\right\}$ – це дотичний (tangent) конус Кларка [3, 4].

Доведення. За припущень 1) та 2) відображення $S(x):B\to R^m$ є однозначним і локально ліпшицевим [5], а також диференційованим за напрямком [6, 7]. Оператор $M:B\to R^p\times R^s_+$, який відображає кожну точку $x_0\in B$ у вектор множників (multipliers) μ_0 та λ_0, також є однозначним і локально ліпшицевим.

За умови строгої доповнюваності $I(x_0,y_0)=J(x_0,y_0)$ для деяких $x_0\in B$, $y_0\in S(x_0)$, коли для кожного активного обмеження $\Psi_k(x_0,y_0)=0$ виконується $\lambda_{0k}>0$, відображення $S(x)$ є диференційованим у точці $x=x_0$ [8]. При цьому градієнт $\nabla S(x_0)$ є оператором, що відображає довільний вектор $u\in R^n$ в єдиний розв'язок задачі квадратичного програмування пошуку інфімума по v функції

$$\frac{1}{2}\left\langle v,Q(x_0)v\right\rangle+\left\langle q(x_0,u),v\right\rangle \quad (11)$$

при обмеженнях

$$\nabla_x \Phi_i(x_0, y_0)u + \nabla_y \Phi_i(x_0, y_0)v = 0, i = 1,..., p, \quad (12)$$

$$\nabla_x \Psi_k(x_0, y_0)u + \nabla_y \Psi_k(x_0, y_0)v = 0,$$

$$k \in I(x_0, y_0), \quad (13)$$

де $q(x_0, u) = \nabla^2_{yx} L(x_0, y_0, \mu_0, \lambda_0)u$.

Коли ж відображення S диференційоване по Гато (Gateaux) в точці $x_0 \in B$, то градієнт $\nabla S(x_0)$ є розв'язком задачі (11)–(13), де замість $I(x_0, y_0)$ розглядається $J(x_0, y_0)$ [9].

За припущень 1), 2) задача дворівневої оптимізації теореми 1 зводиться до пошуку інфімума по $x \in A$ функції $F(x, S(x))$. Оскільки $F(x, y)$ неперервно диференційована, то функція $\Theta(x) = F(x, S(x))$ локально ліпшицева на $B \subseteq A$. Оскільки A – компакт, то існує розв'язок (x^*, y^*) задачі дворівневої оптимізації.

Нехай $\Gamma(x): R^n \to R^r$ – ліпшицевий оператор в околі $x_0 \in R^n$, а σ_Γ – множина точок, де оператор Γ не є диференційованим. Тоді узагальнений якобіан (Jacobian) $\partial \Gamma(x_0)$ оператора Γ у точці x_0 [4] визначається опуклою оболонкою множини матриць $m \times r$

$$conv\left\{ \lim_{t\to\infty} \nabla \Gamma(x_0) : x_t \to x_0; x_t \notin \sigma_\Gamma \right\}.$$

При $r = 1$ такий якобіан називають узагальненим градієнтом (субдиференціалом) Γ у точці x_0.

У роботі [4] доведено, що точці $x_0 \in A$

$$\partial \Theta(x_0) = \nabla_x F(x_0, S(x_0)) + \quad (14)$$

$$+ \{[P(x_0)]^T \nabla_y F(x_0, S(x_0)) P(x_0) : P(x_0) \in \partial S(x_0)\}.$$

Лема 1. Нехай C – симетрична позитивно визначена матриця $m \times m$, H – матриця $l \times m$ рангу l, а E – відображення $w^* \in R^m$ в єдиний множник Лагранжа задачі квадратичного програмування

$$\frac{1}{2}\langle z, Cz \rangle - \langle w^*, z \rangle \to \inf_z, \quad (15)$$

$$Hz = \vec{0}. \quad (16)$$

Тоді приєднане відображення E^T кожному $v^* \in R^l$ ставить у відповідність єдиний розв'язок задачі квадратичного програмування

$$\frac{1}{2}\langle w, Cw \rangle \to \inf_w, \quad (17)$$

$$Hw = v^*. \quad (18)$$

Доведення леми. Необхідними й достатніми умовами розв'язку для задачі (15), (16) є

$$\begin{bmatrix} C & H^T \\ H & \vec{0} \end{bmatrix} \begin{bmatrix} z \\ v \end{bmatrix} = \begin{bmatrix} w^* \\ \vec{0} \end{bmatrix},$$

а для задачі (17), (18) –

$$\begin{bmatrix} C & H^T \\ H & \vec{0} \end{bmatrix} \begin{bmatrix} w \\ \rho \end{bmatrix} = \begin{bmatrix} \vec{0} \\ v^* \end{bmatrix},$$

де v – множник Лагранжа задачі (15), (16).

Звідси розв'язок задачі (17), (18) однозначно визначається через $w^* \in R^m$:

$$v = G^{-1} H C^{-1} w^*,$$

де $G = HC^{-1}H^T$ – матриця Грама (Gramm) у C-метриці. Розв'язок задачі (15), (16) однозначно визначається через $v^* \in R^l$:

$$w = C^{-1} H^T G^{-1} v^*.$$

Продовження доведення. Враховуючи вираз (14) для субдиференціала та умови оптимальності [10], достатньо показати

$$[\nabla S(x^*)]^T \nabla_y F(x^*, y^*) =$$

$$= -[\nabla^2_{yx} L(x^*, y^*, \mu^*, \lambda^*)]^T t - [D(x^*)]^T \eta.$$

За умови сторогої доповнюваності градієнт $\nabla S(x^*)$ визначається як розв'язок задачі квадратичного програмування (11)–(13). Умова Куна–Такера для цієї задачі при $x_0 = x^*$, $y_0 = y^*$

$$Q(x^*)v + q(x^*, z) + [C(x^*)]^T \xi = 0, \quad (19)$$

де: $\xi \in R^{p+o}$ – єдиний множник, що відповідає обмеженню-рівності

$$C(x^*)v = -D(x^*)z; \quad (20)$$

o – кількість активних обмежень $\Psi_j(x^*, y^*) = 0$.

Оскільки за умовами теореми $Q(x^*)$ – позитивно визначена матриця, то з (19) випливає

$$v = -[Q(x^*)]^{-1}\{q(x^*, z) + [C(x^*)]^T \xi\}, \quad (21)$$

звідки в силу (20) маємо

$$C(x^*)[Q(x^*)]^{-1}\{q(x^*,z)+[C(x^*)]^T\xi\}=D(x^*)z,$$

$$G(x^*)\xi=C(x^*)z-C(x^*)[Q(x^*)]^{-1}q(x^*,z),\quad(22)$$

де $G(x^*)=C(x^*)[Q(x^*)]^{-1}[C(x^*)]^{-1}$ – матриця Грамма у $Q(x^*)$-метриці. Оскільки ця матриця позитивно визначена, то з (21), (22) випливає

$$v=v_1+v_2,$$

$$v_2=-[Q(x^*)]^{-1}[C(x^*)]^T[G(x^*)]^{-1}D(x^*)z,$$

$$v_1=[Q(x^*)]^{-1}\{-q(x^*,z)+$$

$$+[C(x^*)]^T[G(x^*)]^{-1}C(x^*)[Q(x^*)]^{-1}q(x^*,z)\}.$$

Отже, розв'язок задачі (11)–(13) має вид

$$v=\Xi_1[-\nabla^2_{yx}L(x^*,y^*,\mu^*,\lambda^*)]+\Xi_2[-D(x^*)],\quad(23)$$

де: Ξ_1 – симетричний проектор $[Q(x^*)]^{-1}b$, $b\in R^m$, на систему лінійних рівнянь $\{v\in R^m : C(x^*)v=\vec{0}\}$ у $Q(x^*)$-метриці; Ξ_2 – відображення $a\in R^{p+o}$ на вектор найменшої норми у $Q(x^*)$-метриці, що задовольняє системі лінійних рівнянь $\{v\in R^m : C(x^*)v=a\}$. Тоді в силу співвідношення (23)

$$\left\langle\nabla_y F(x^*,y^*),vz\right\rangle=$$

$$\left\langle\nabla_y F(x^*,y^*),\Xi_1[-\nabla^2_{yx}L(x^*,y^*,\mu^*,\lambda^*)]z+\right.$$

$$\left.+\Xi_2[-C(x^*)]z\right\rangle=$$

$$=\left\langle[-L(x^*,y^*,\mu^*,\lambda^*]^T\,\Xi_1\nabla_y F(x^*,y^*),z\right\rangle+$$

$$+\left\langle[-D(x^*)]^T(\Xi_2)^T\nabla_y F(x^*,y^*),z\right\rangle,$$

де за лемою 1

$$\Xi_1\nabla_y F(x^*,y^*)=t,\quad [\Xi_2]^T\nabla_y F(x^*,y^*)=\eta.$$

Задача (1)–(6) має структуру задач [11] з умовами оптимальності [12].

Список використаних джерел

1. *Suh S., Kim T.J.* Solving nonlinear bilevel programming models of the equilibrium network design problem: a comparative review // Annals of operations research. – 1992. – 34. – P. 203–218.
2. *Wardrop J.G.* Some theoretical aspects of road traffic research // Proceedings of Institute of Civil Engineers. – 1952. – 1. – P. 325–328.
3. *Outrata J.V.* Necessary optimality conditions for Stackelberg problems // Journal of optimization theory and applications. – 1993. V. 76. – P. 305–320.
4. *Кларк Ф.* Оптимизация и негладкий анализ. – М.: Наука, 1988. – 280 с.
5. *Cornet B.*, *Laroque G.* Lipschitz properties of solutions in mathematical programming // Journal of optimization theory and applications. – 1987. – V. 53. – P. 407–427.
6. *Jitornrum K.* Solution point differentiability without strict complementarity in nonlinear programming // Mathematical programming study. – 1984. – V. 21. – P. 127–138.
7. *Kyparisis J.* Sensitivity analysis for nonlinear programs and variational inequalities with nonunique multipliers // Mathematics of operations research. – 1990. – V. 15. – P. 286–289.
8. *Fiacco A.V.* Sensitivity analysis for nonlinear programming using penalty methods // Mathematical programming. – 1976. – V. 10. – P. 287–311.
9. *Malanowski K.* Differentiability with respect to parameters of solutions to convex programming problems // Ibid. – 1985. – V. 33. – P. 352–361.
10. *Hiriart-Urruty J.B.* Tangent cones, generalized gradients and mathematical programming in Banach spaces // Mathematics of operations research. – 1979. – V. 4. – P. 79–97.
11. *Горбачук В.М., Гаркуша Н.І.* Максимізація функції доходу та мінімізація функції витрат у дворівневому програмуванні // Вісник Київського університету. Серія: фізико-математичні науки. – 2005. – № 4. – С. 147–152.
12. *Горбачук В.М., Гаркуша Н.І.* Задачі дворівневого прийняття рішень // PDMU-2006 (21–25 травня 2006 р., Східниця, Україна). – К.: КНУ ім. Т. Шевченка, 2006. – С. 178–180.

Надійшла до редколегії 15.04.2010